

TACKLING FOOTBALL HOOLIGANISM

A Quantitative Study of Public Order, Policing and Crowd Psychology

Clifford Stott
University of Liverpool

Otto Adang
University of Liverpool and
Police Academy of the Netherlands

Andrew Livingstone
Cardiff University

Martina Schreiber
University of Liverpool and
Police Academy of the Netherlands

This paper contributes to the science of crowd dynamics and psychology by examining the social psychological processes related to the relative absence of “hooliganism” at the Finals of the 2004 Union Européenne de Football Association (UEFA) Football (Soccer) Championships in Portugal. Quantitative data from a structured observational study is integrated with data from a questionnaire survey of a group associated ubiquitously with ‘hooliganism’ – namely England fans. This analysis provides support for the contention that the absence of ‘disorder’ can be attributed in large part to the non-paramilitary policing style adopted in cities hosting tournament matches. Evidence is presented which suggests that this style of policing supported forms of non-violent collective psychology that, in turn, served to psychologically marginalise violent groups from the wider community of fans. The study highlights the mutually constructive relationships that can be created between psychological theory, research, policing policy and practice, particularly in relation to the successful management of ‘public order’. The paper concludes by exploring some of the wider implications of this research for theory, policy, the management of crowds, social conflict, and human rights more generally.

Keywords: policing, crowds, hooliganism, social identity

Association football or soccer is arguably the most popular spectator sport in

Clifford Stott, School of Psychology, University of Liverpool, Liverpool, United Kingdom; Otto Adang, School of Psychology, University of Liverpool, Liverpool, United Kingdom, and Police Academy of the Netherlands, Apeldoorn, the Netherlands; Andrew Livingstone, School of Psychology, Cardiff University, Cardiff, Wales; Martina Schreiber, School of Psychology, University of Liverpool, Liverpool, United Kingdom, and Police Academy of the Netherlands, Apeldoorn, the Netherlands.

This research report was made possible by grants from the Economic and Social Research Council (RES-000-23-0617) and from the U. K. Home Office. Data collection was made possible through the long-term co-operation of the U. K. Home Office, especially David Bohannon and Martin Goodhay; the Portuguese Public Security Police, especially José Leitão, Paulo Pereira, José Neto, Luis Simões, and João Pires; the British Consulate in Portugal, especially Gary Fisher and Glynne Evans; the Euro2004 Organising Committee, especially Paulo Gomes and Luís Trindade Santos; the Football Supporters Federation, especially Kevin Miles; and the numerous England fans that have supported our research. Thanks should also be extended to the U. K. Police Euro2004 delegation, especially David Swift; and the Football Association, especially Ian Murphy, Nicola Jones, and Doug Hopkins.

Correspondence concerning this article should be addressed to Clifford Stott, The Henri Tajfel Laboratory, School of Psychology, University of Liverpool, L69 7ZA. U. K. E-mail: c.stott@liverpool.ac.uk

the world. The game has historically attracted large crowds who gather on a regular basis in sometimes massive stadiums across South America, Europe, Africa, most of Asia, the Middle East, and more recently the USA. But this global crowd phenomenon also carries with it an historical and ubiquitous association with major incidents of public disorder. For example, Dunning (2000) reports upon incidents of football crowd disorder throughout the twentieth century in 37 different countries as diverse as China, India, the USA, Argentina, and Ireland.

Such incidents can and do have disastrous consequences. In Europe, the 1985 Heysel stadium disaster occurred when rioting involving fans of English and Italian club sides led to the collapse of a defective wall and the subsequent death of 39 mostly Italian fans. While there was never an official enquiry into the disaster, it was understood by leading policy makers and theorists to have been caused by English hooligans (Poplewell, 1986; Williams et al., 1989). As a direct response to Heysel, the Union Européenne de Football Association (UEFA) banned all English football club sides, and therefore their fans, from European Competition for a period of five years.¹ As Dunning argues, it was Heysel, more than any other event, that “fixed the idea of football hooliganism as an ‘English disease’ firmly in the minds of people around the world” (2000).

Thus one of the largest and most politically significant incidents of football crowd disorder, one that has formed the background context for contemporary understanding and international policy, occurred at a football match between teams from different nation states. Indeed, the largest stadium tragedy ever recorded also occurred during an international football fixture, at the Lima Stadium between Peru and Argentina in 1964.² Here as many as 328 fans died and approximately 5000 were injured after police reportedly closed exit gates and fired tear gas into the crowd in an attempt to quell violent disturbances over a disputed goal. In this sense, it is relevant to address such tragedies not so much in terms of hooliganism but more precisely as complex crowd events with an international dimension.

Given that they are crowd events, football matches provide important opportunities for empirical study and development in theoretical understanding of crowd dynamics (Stott, Adang, Livingstone, & Schreiber, 2007; Stott, Hutchison, & Drury, 2001; Stott & Reicher, 1998a) and public order policing (Adang, 1991; Della Porta & Reiter, 1998; Jefferson, 1987, 1990; Stott, Livingstone, & Hoggett, 2008; Waddington, 1987, 1991, 1993, 1994). Additionally, in the attempts to control incidents of football crowd disorder, laws have been created that arguably undermine fundamental civil liberties and human rights. For example, in the U.K., legislation has been enacted that makes it possible for the authorities to place severe restrictions upon individual freedom of movement for suspected hooligans, even in the absence of evidence that would be sufficient to achieve a criminal conviction. Civil liberties groups within the U.K. have expressed serious concerns

¹ Liverpool F. C. was banned for a further one year.

² The Kuzhniki stadium disaster claims an unofficial death toll of 340. This disaster took place at the Lenin Stadium in Moscow during a UEFA Cup match between F. C. Spartak Moscow (Russia) and HFC Haarlem (Netherlands).

that this legislation is illegal because it contravenes one of the key principles of the rule of law (for fuller discussion of these issues, see Stott & Pearson, 2006, 2007).

Despite its widespread acceptance the concept of football hooliganism is considered by some researchers as largely a mass media construction (Crabbe, 2003; Poulton, 2005; Stott & Pearson, 2007). It is problematic as a scientific and technical construct primarily because it is used to describe such a vast array of different forms of activity, ranging from individual drunkenness to spontaneous rioting involving hundreds of supporters. Nonetheless, there is wealth of academic theory and research on the causes of hooliganism (e.g., Armstrong, 1998; Dunning, Murphy, & Williams, 1988; Kerr, 1994; Marsh, Rosser, & Harre, 1978). While there is little consensus within this literature, the different schools of thought tend to share the assumption that football crowd disorder is caused by the convergence of fans who have some form of predisposition toward violent confrontation (see Frosdick & Marsh, 2005, for an overview). Moreover, theory in this area is primarily focused upon explaining the origins of hooliganism within different nation states.³ As a consequence this body of research has done little to explore empirically how the dynamics of football crowd events can and do actually function (except see King, 1995). Arguably then, this body of theory has limited relevance for debates about the dynamics of crowd disorder at matches and tournaments with an international dimension.

The low levels of football crowd disorder at the 2004 UEFA European Championships Finals in Portugal (Euro2004) have led it to be regarded, in security terms at least, as one of the most successful international football tournaments ever staged in Europe (Home Office, 2005). This is particularly striking given the presence in Portugal during the tournament of an estimated 150,000 England fans (Independent Football Commission, 2004), a fan group that has a history of involvement in major riots. Indeed, Dunning (2000) acknowledges that by 1998 many commentators understood that “whenever English supporters are taking part in an international tournament, it is inevitable there will be trouble” (p. 149). The serious rioting involving English fans at the World Championships in France that year (France98) apparently proved such sentiments correct (but see Stott et al., 2001). Indeed, at the 2000 European Championships (Euro2000), co-hosted by Belgium and the Netherlands, there were also a series of major incidents of disorder during which 965 England fans were arrested. At the 2006 World Cup Finals in Germany (WM2006), there were again a series of incidents of crowd disorder and 828 England fans were arrested or detained (for a detailed overview of these events see Stott & Pearson, 2007). Even at the 1996 European Championships in England (Euro96), there were a series of incidents of football related disorder involving English people, both within and outside the country. However, during Euro2004, the Portuguese and British authorities recorded no major incidents of crowd disorder in cities hosting matches (Home

³ Within England – Armstrong (1998); Dunning et al. (1988); Marsh et al. (1979); Italy – Dal Lago (1990) and Roversi (1991); Germany – Pilz (1996); the Netherlands – van der Brug (1989) and Russell and Goldstein (1995).

Office, 2005).⁴ What is also particularly striking, in contrast to previous tournaments, is that in these cities only one England fan was arrested for a violence-related offence.

Nonetheless, arrest figures along with other official and unofficial indicators, such as reports of assaults, damage to property, and media reports, are notoriously unreliable and problematic as a scientific measure of the levels of public disorder during crowd events. There can be biases introduced by the political climate, the way reports are written, how crimes are registered, or in the way crimes are categorised, even by official sources. Moreover, arrest figures can be as much a reflection of police tactics of preventative mass arrest as much as any reflection of the overall levels of disorder. For example, of the 965 England fans arrested during Euro2000, only one was subsequently convicted of any criminal offence. Nonetheless, although arrest figures are an imperfect measure of disorder, Euro2004 is widely recognised as standing apart from other major international football tournaments held in mainland Europe (Home Office, 2005; Independent Football Commission, 2005).

The issue of how such low levels of disorder actually occurred is a fundamental one for theory, policy, and practice. Yet, it is difficult to understand such variation in terms of dominant theories of football hooliganism, unless one also assumes that the relative harmony was because those predisposed toward disorder (i.e., hooligans) were unable to attend. Therefore, one prominent explanation for the positive outcome is that it was the result of legislation, most notably in the U.K. and Germany, which prevented “known hooligans” from leaving their home country during the tournament (Home Office, 2005).⁵ In other words, positive outcomes were achieved through controlling the movement across national borders of those individuals and groups who were seen by the authorities as likely to initiate disorder. We see this explanation as problematic for a number of reasons, not least of all because it is unclear how it can account for both the absence of crowd disorder at Euro2004 and the re-emergence of disorder at WM2006.

This paper provides a quantitative empirical assessment of an alternative explanation; that the “public order” successes of Euro2004 were in large part due to the relationships developed between research on crowd psychology and police public order management policy and practice in match cities (see also Stott, Adang, Livingstone, & Schreiber, 2007; Stott & Pearson, 2007). In doing so, the paper contributes to a body of evidence that examines the policy and theory implications of this alternative explanation and highlights the mutually construc-

⁴ It is worth noting that the Independent Football Commission (2005) does report rioting involving up to 500 German and Dutch fans in Oporto, but no other sources of data (e.g., semi-structured and structured observations; interviews with senior commanders in Oporto – see Stott et al. (2007) – German, Dutch, and Portuguese police; government and fan organisation statements and reports; all media including television, press, and web articles) ever record such an incident taking place. Given this overwhelming consensus, it is safer to assume that the IFC report is inaccurate in this respect.

⁵ UEFA made it clear that if there were significant incidents of disorder involving England fans the national team would be expelled from the tournament. This could have contributed to an increased awareness among England fans of the importance of not becoming involved in disorder. However, this was never referred to as an issue by fans in the open-ended items within the survey used in this study.

tive relationship that can be created between social psychological theory and the effective management of public order.

Crowd Dynamics, Social Identity, and Intergroup Interaction

In contrast to classic models of crowd psychology (Allport, 1924; Le Bon, 1895/1947; Zimbardo, 1969) but in line with others (e.g., McPhail, 1991; Turner & Killian, 1987), the Elaborated Social Identity Model of crowd behaviour (ESIM) posits that collective behaviour during crowd events is fundamentally normative. According to the ESIM these norms are an outcome of crowd participants' shared and salient social identity. This social identity is represented at the level of the individual in terms of cognitive social categories. These category memberships form part of the self-concept and represent context-dependent social judgements that are dynamic in both form and content (Turner, Oakes, Haslam, & McGarty, 1994).

The ESIM proposes that crowd members' social identity is embedded within intergroup relations that are defined in terms of legitimacy and power (cf. Tajfel & Turner, 1979). Social identity is defined here as a cognitive model (or perception) of the legitimacy of one's shared position within such social contexts along with the collective actions that are possible and legitimate given such a position (Drury & Reicher, 2000, p.581). It follows from this that the form (i.e., who does or does not share that position) and content of crowd participants' social identity (i.e., what is understood as legitimate behaviour) can radically alter because of the way in which the actions of the police affect the intergroup relationships during a crowd event.

The ESIM also proposes that changes in the form and content of crowd members' social identity lead to changes in (a) who or what behaviours (and ideas) can become influential within the crowd, and (b) the forms of collective action that can and do subsequently emerge. Thus, while the ESIM accepts that crowd members' social identity is a meaningful *reflection* of intergroup relations, it also emphasises that acting collectively on the basis of that social identity can in turn *change* the intergroup context, in particular for the police. The ESIM therefore proposes that to understand the dynamics of crowd events, it is necessary to conceptualise and study crowd psychology as an ongoing intergroup *process* of which the police can be an integral component.

The ESIM has been developed and validated through a series of studies of riots in a variety of different settings, including anti-tax (Drury & Reicher, 1999; Stott & Drury, 2000), environmental (Drury & Reicher, 2000) and student (Reicher, 1996) protests within the U.K.. Of particular relevance here is the utility of the theory in accounting for both the presence and absence of rioting amongst British football fans travelling to international football tournaments (Stott et al., 2007; Stott et al., 2001; Stott & Reicher, 1998a). For example, Stott and Reicher (1998a) provide a qualitative analysis of the social psychological processes through which rioting between fans of the England national football team and police developed during the 1990 World Cup Finals in Italy. One key element of this analysis was to highlight the existence of asymmetries in how Italian police and England fans understood the same intergroup context. Whereas "ordinary" England fans tended to see themselves as differentiated from hooligans and

engaging in normatively appropriate behaviour (e.g., singing loudly in large groups), their experiences suggested that they were viewed and treated by police as *uniformly* dangerous hooligans.

A second key element of Stott and Reicher's (1998a) analysis was to suggest that these asymmetries were important because of the ability of one group (i.e., the police) to then *impose* their understanding of the context onto the other group (i.e., England fans) through relatively undifferentiated forms of coercive police intervention (e.g., baton charges). This intergroup interaction in turn served to unite large numbers of ordinary England fans and hooligans around a common understanding of "victimhood" and an emergent perception of the legitimacy of violent "retaliation" against the police. Moreover, this emergent collective identity also *empowered* England fans, such that retaliation was not only seen as *proper* but also as *possible* social action. This retaliation precipitated an upward spiral of conflict, culminating in a large scale riot. In other words, police expectations of a uniformly violent group acted as a self-fulfilling prophecy because of the effect of their subsequent strategy and tactics on fans' social identity.

In addition to these analyses of processes within crowds, the ESIM has also developed through analyses of police understanding and strategy in relation to crowd events. These have highlighted how police orientations can be shaped by an understanding of crowds as holding an inherent potential for violence (based largely upon classic models of the crowd, e.g., Le Bon, 1895/ 1947). This understanding is then used to inform and legitimise the kinds of undifferentiated forms of coercive police intervention that literature in crowd psychology (Drury, Stott, & Farsides, 2003; Stott & Drury, 2000; Stott & Reicher, 1998b) and public order policing (Della Porta & Reiter, 1998; Jefferson, 1987; Waddington, 2005) implicates in the escalation of disorder. However, while these strands of research on crowds and police officers clearly tally in terms of the ESIM's theoretical model of crowd dynamics, they have not previously been combined within a single study of the same event (Reicher, 1996; Stott & Reicher, 1998b). Moreover, previous ESIM studies of crowd events have all been based upon qualitative data and have therefore been unable to provide quantitative measurements of police deployments or collective psychology and behaviour as they relate to specific crowd events.

Constructing Relationships Between Theory and Practice

Nevertheless, the ability of the ESIM to provide a powerful and parsimonious explanation of crowd dynamics in a range of settings, both violent and non-violent, suggests its considerable relevance for public order policing and legal policy. This is particularly so in light of other debates about the association between policing and crowd behaviour (Adang, 1988, 1991, 1998; Baker, 2005; Della Porta & Reiter, 1998; Hall & De Lint, 2003; Innes, 2005; Jefferson, 1987; King & Waddington, 2004, 2005; Sheptycki, 2005; D. Waddington, 2005; Waddington & King, 2005; Waddington, 1987, 1991, 1993, 1994). More specifically, a large scale structured observational study of police deployment at Euro2000 identified two distinct and contrasting forms of policing. These were characterised as low versus high profile (that relate loosely to the descriptions of "negotiated management" versus "escalated force" outlined in Waddington, 2005, p. 10), where high profile deployments were defined relative to low profile policing in

terms of the much greater visibility of uniformed police, riot police (i.e., officers wearing protective equipment), and police riot vehicles (i.e., vans and water cannon). Observations in high profile cities also identified a relative absence of low impact intervention tactics combined with less and poorer quality of communication between police and fans prior to the emergence of disorder. The study also demonstrated that the greatest levels of disorder occurred during events defined by the police as presenting a low risk of disorder but where high profile policing had been utilised. In contrast, during events defined as high risk there were no significant differences in the levels of observed disorder between cities utilising high and low profiles (Adang & Cuvelier, 2001).

As well as reliably quantifying both police deployment and levels of disorder this study was important because it exposes the problems policy makers face in creating accurate risk assessments of threats to public order at football events with an international dimension. They also powerfully demonstrate, once again, the potentially counterproductive impact of high profile policing and the relative efficiency of the low profile approach (see Waddington, 2005, for an overview of research in this area). Moreover, these data support the contention that the general patterns of collective action observed during such tournaments can be related meaningfully to the nature of interactions between police and fans. Nonetheless, just as ESIM's preoccupation with underlying psychology has so far meant that a quantitative analysis of police deployment and disorder has not been forthcoming, Adang and Cuvelier's (2001) focus on the quantitative analysis of broad patterns of behaviour meant that the psychological processes underlying these patterns could not be examined.

These strands of research were subsequently brought together through research, funded by the U.K. Home Office, focused upon the development of a model of good practice for policing high-risk crowds of foreign nationals attending football matches with an international dimension. In summary, this model proposes that disorder can be minimized during such events by avoiding disproportionate and indiscriminately forceful intervention against crowds. In order to achieve this, police should hold a strategic orientation toward the facilitation of crowds and the tactical profile should be information-led, graded, and dynamic. Therefore, policing should be initially non-confrontational (e.g., avoiding the use of paramilitary uniforms) and focused on monitoring and information gathering. If problems are subsequently identified, police use of force must be information-led to increase the likelihood that it can be targeted specifically against those transgressing previously defined and communicated behavioural limits. Having dealt with the risk, the police should also have a clearly defined policy of de-escalation (see Stott & Adang, 2003a, 2003b; Stott & Pearson, 2007, pp. 233-255, for detailed overviews of this model; see also Police Cooperation Working Party of the Council of the European Union, 2006).

Research, Policy, and Practice at Euro2004

In 2001, a group of senior officers from the PSP (Policia de Seguranca Publica – one of Portugals two main police forces), charged with the development of policy on the use of force for Euro2004, became aware of the potential relevance of this research. In particular, they sought to utilize it to support the development of low profile policing for the tournament in opposition to pressures they faced within

Portugal to develop a high profile approach. This collaboration not only supported the development of their policy but also subsequently allowed the current authors to conduct a major scientific evaluation of its implementation.⁶

This evaluation process has already provided a qualitative analysis of policing and of England fans collective phenomenology and behaviour during the tournament (Stott et al., 2007). That study suggests that the PSP did implement a low profile approach that was in line with the model of good practice and that this profile was associated with low levels of disorder. Additionally, Stott et al. (2007) suggest that England fans consequently perceived their intergroup relations with the PSP as legitimate. In this intergroup context, England fans developed bonds of common identification with football fans from other nations and differentiated themselves psychologically from hooligans. It is argued that this created a context in which violent fans were less able to influence crowds and thus remained marginalised and disempowered. Moreover, the study introduces evidence of self-policing or norm enforcement among England fans, such that attempts to create disorder were actively suppressed by other fans.

The Current Paper

The analysis presented by Stott et al. (2007) offer further support for the ESIM. They also provide substantial evidence of the effectiveness of policing policy developed with an awareness of this theoretical approach (Reicher, Stott, Cronin, & Adang, 2004; Reicher et al., 2007). However, the existing research is limited by its almost exclusive focus upon *qualitative* analyses of both police deployment and fans collective perceptions and behaviour. Thus, these studies lack any systematic measurement of critically important facets proposed by the ESIM. Moreover, existing analyses of crowd events are limited by the fact they focus independently upon one side or other of the intergroup dynamic (i.e., either crowd members *or* police).

The aim of the present paper is to address these limitations by presenting an integrated *quantitative* analysis of the policing of and perceptions and behaviour among crowds of England fans attending Euro2004. In other words, we aim to provide an integrated study of the two sides of the intergroup dynamic as this relates to the same crowd events. In doing so, we explore the implications of the ESIM for policing policy toward crowds, both at Euro2004 and more generally. We therefore begin with an analysis of data gathered through a programme of structured observations of policing deployment and crowd behaviour. These data on policing are then supplemented by a survey-based study of England fans perceptions.

Method

Procedure for Structured Observations

Structured observations of fourteen matches were conducted between June 12 and July 1, 2004. The method directly mirrored that employed by Adang & Couvélér (2001). This was done in order to allow statistical comparison between

⁶ This evaluation was made possible through a grant from the Economic and Social Research Council (grant ref: RES-000-23-0617).

the current data and a data set obtained from Euro2000 (for a full breakdown of the Euro2000 method, see Adang & Cuvelier, 2001). Observations for the present study were conducted around all matches involving the national teams of Portugal, the Netherlands, Germany, and England. This selection ensured that an equal number of matches were sampled that had been defined by the authorities as posing increased and normal risk to public order.⁷

Observations were conducted by 16 bilingual Portuguese nationals. Half were Psychology postgraduates or final year undergraduates recruited from universities in Portugal, and half were final year police cadets recruited from the Instituto Superior de Ciências Policiais e Segurança Interna (the Portuguese National Police Academy) in Lisbon. The observation teams were trained at a three-day workshop preceding the tournament. This included lectures on theoretical background, the observation protocol, observation techniques, and observation exercises. The workshop culminated in a test observation around the Portuguese Cup Final on May 16, 2004. Supervision by the first two authors during this test observation took place in order to ensure the observers were sampling according to the protocol. The literature does suggest that this method combines ease of scoring with high inter-observer reliability (Lehner, 1979). A subsequent debriefing with all observers was conducted to assess the reliability of inter-observer agreement, which was assessed in terms of qualitative consensus.

During the tournament, the observers were divided into four teams; two based in the south and two in the north of Portugal, each team containing equal numbers of students and police cadets. Assignment of teams to matches was determined prior to the tournament based upon their availability, equal sharing of workload, and the constraints of the tournament schedule, allowing the possibility for threeday work periods: one day preparation and evening observation, one match day observation, one day working on field notes and data input, followed by at least one rest day.

Each observation was carried out by one team operating as two pairs, with each pair consisting of one student and one cadet. This mixing of students was done deliberately to minimise observer bias. Observations generally started at 20.00 hours on the day preceding the selected match and continued until 00.00 hours before recommencing at 12.00 hours on the match day and finishing twelve hours later at 00.00. Each pair was instructed to gather in separate locations where large crowds of fans had gathered (e.g., at official fan zones, in public squares, outside of stadiums). Observers were instructed to take their distance whenever feeling unsafe, but none expressed concerns that they felt threatened at any point and all walked freely between fan groups to position themselves in such a way as to best sample the data required. Each observer within each pair selected a separate physical area within that geographical location and began observations as instructed for at least one hour, after which they could move to a different location.

Every fifteen minutes, observers took note of a series of predefined categories.

⁷ The Portuguese authorities had three classifications of risk: normal, medium and high. We have collapsed high and medium into a single increased risk category. All matches played by Germany and England had been classified as high or increased risk.

These included overall numbers of fans and police, the nature of police deployment, and the quality and quantity of interactions between fans and police. The method for this aspect of the observation is referred to as scan sampling, a special form of instantaneous sampling also referred to as “point” (Dunbar, 1976) or “on the dot” (Slater, 1978) sampling. This technique is effective for measuring behavioural states and the behavioural categories were clearly delineated to assist in quick sampling. The fifteen-minute time period was selected in order to allow at least four samples per hour while allowing enough time for each sample to be fully and accurately recorded.

At each sample time, observers were also required to indicate if an “incident” had taken place in the preceding 15 minutes and if it was small (a quarrel involving at most four individuals), medium (a quarrel involving at least five and at most 10 individuals), or large (a quarrel involving more than 10 individuals). They were also required to record whether police intervention with use of force had taken place. Thus, this aspect of the observation adopted a one-zero sampling technique. This method is often referred to in the literature as “time sampling” (Hutt & Hutt, 1974). Again, the fifteen-minute time period was felt to be optimal, given that police interventions and use of force are relatively rare. Specifically, shorter time intervals might have yielded more samples, but it is most unlikely that they would have provided more in the way of meaningful data. Moreover, as incidents that contain more than one act last on average three minutes (Adang, 1991), the time interval chosen should not be too close to this figure to decrease the likelihood that the same event occurs in two samples. The fifteen-minute time sampling period was, therefore a useful means of balancing the competing requirements of sample volume and sample quality.

Given that each observer was observing a separate physical location, there was no basis upon which measures of inter-observer reliability could be calculated. Other steps were therefore taken to maximise the reliability of the samples. Specifically, the first, second, and fourth authors supervised each of the observation teams, meeting with them on a regular basis throughout each observation period, checking for any irregularities in the recorded data. Data from the structured observations were recorded directly onto a paper spreadsheet and a complete breakdown of the categories employed in the structured observations is included in Appendix 1. A total of 1896 observational samples were taken. These are subdivided into seven increased risk matches (899 samples) and seven normal risk matches (997 samples).

Fan Survey

Data were collected from fans via web-based questionnaires located at URL www.footballfans.org.uk. The pre-tournament questionnaire was posted two weeks before and removed on the day the tournament began. Immediately following each of Englands four matches, a questionnaire relating to the host city (i.e., Coimbra or Lisbon) was posted and remained active until one month after the tournament had concluded. The pre- and post-tournament questionnaires began by asking fans age, gender, and how many times they had previously travelled abroad to support England. Both questionnaires included a five-item scale ($\alpha = .77$) measuring ingroup (i.e., England fan) identification adapted from Doosje, Elle-

mers, and Spears' (1995) and Haslam, Oakes, Reynolds, and Turner's (1999) four- and one-item measures. These were followed by a series of items designed to measure key variables. These were perceived similarity with other groups, differentiation from violent and others, expectations about – and experiences of – intergroup relations. These are described in the pre- and post-tournament measures subsections below.

Except for age, gender, and match-going history, fans responded to all items by clicking check boxes on five-point Likert-type scales ranging from *strongly agree* to *strongly disagree*. Both questionnaires also included two open-ended items for fans to provide their qualitative impressions (see Appendix 2).

Pre-tournament measures. The five ingroup identification items were followed by one item measuring the extent to which fans differentiated themselves from violent others, and one item measuring fans perceived similarity with opposition fans. One item also measured fans perceived similarity to the Portuguese Police in match cities. High scores on these items represent high levels of identification, differentiation, and similarity, respectively.

In order to measure fans expectations of intergroup relationships, two items gauged expectations about the local population in match cities ($r = .63, p < .001$), and three items gauged expectations about the police in match cities ($\alpha = .88$). Finally, one item gauged expectations about relations with opposition fans. High scores on these items represent positive expectations.

Post match measures. The post-match questionnaires were based upon the pre-tournament measures but included additional items. There were no significant differences on any items across all four of the post-tournament questionnaires relating to the different matches. Therefore, the data were combined into a single post-tournament data file.

The five ingroup identification items ($\alpha = .76$) were again followed by single item measures of differentiation from violent others, perceived similarity to opposition fans, and perceived similarity to the Police in match cities. High scores on these measures represent high levels of identification, differentiation, and similarity, respectively. Five-item scales gauged fans experiences of their ingroup ($\alpha = .90$) and fans experiences of the local population in match cities ($\alpha = .90$). A single item gauged fans experiences with opposition fans, while a six-item scale ($\alpha = .93$) gauged fans experiences of policing in match cities.

Participants. England fans were recruited in a number of different ways: through small articles on the Football Associations (F.A.) official "EnglandFans" website; in five editions of *Freelions* (a magazine produced by an independent football fan organisation called the Football Supporters Federation); through email distribution loops set up by other independent fan organizations; and through flyers advertising our research and website URL that were distributed by the authors during the tournament. In addition, before the tournament, the first author attended seven meetings held by England fan organisations. At these meetings, fans were encouraged to subsequently complete both the pre- and post tournament online questionnaires.

One hundred and two questionnaires were submitted, 39 pre-tournament and 63 within one month of the end of the tournament. The relatively low response rate is typical for football fans who tend to lack motivation to engage in research, particularly when events have been positive for them. Three of the pre-tournament

respondents were female. The mean age was 34 years; 18% had never travelled abroad as an England fan before; and 33% had travelled between one and five times previously. The remaining 49% had travelled abroad with England six or more times. Thirteen of the post-tournament respondents were female and four did not specify their gender. The mean age was 35 years; 73% had travelled abroad to watch England between one and five times previously, with the remainder having travelled abroad with England more than six times. We were not able to check that those completing the questionnaires had actually travelled to the events in question. Our assumption is that the recruitment technique and wording of items was such that it would have been unlikely that fans would have known about and therefore completed the questionnaires without having travelled to the host cities for England's matches.

Results

Policing and Public Order

The data from the structured observations indicate that there was a visible police presence in only 56 % of samples. There was a significant difference in levels of visible police deployment ($\chi^2 = 6.22, p < .05$) between normal and increased risk events. This difference was not in the expected direction, with increased risk situations recording a 53 % visible police presence and normal risk showing a 59 % police presence. On closer analysis, this difference is entirely due to the special case of the match between England and Switzerland in Coimbra, which was increased risk but which showed only a 19 % police visibility. Excluding Coimbra, the difference is non-significant ($\chi^2 = 0.22, p > .05$). Across both normal and increased risk situations, 42 % of samples saw no visible police presence during crowd events. This overall level of police visibility is significantly lower than Euro2000 ($\chi^2 = 205.5, p < .001$).

If police were visibly present, the proportion was on average 5.5 officers per 100 fans, with on average 4.5 officers per 100 fans in relation to normal risk matches and 6.9 officers per 100 fans around increased risk matches ($U = 518, z = 2.357, p < .05$). This compares to a visible presence during Euro2000, which ranged from an average of 6 officers per 100 fans in low-profile, normal risk situations to 50 officers per 100 fans in high-profile high-risk situations.

If uniformed police were visibly present at sample time, this was never in the form of the Corpo de Intervenção (the PSP's riot police) deployed wearing their full protective equipment (i.e., 0% of samples). During Euro2000, fully equipped riot police were present in 15 % of samples, with increased risk situations having on average more than twice as many officers deployed wearing protective equipment than normal risk situations (Adang & Cuvelier, 2001). During Euro2004, police were present in partial riot gear (i.e., Corpo de Intervenção or other police deployed with but not wearing protective equipment) in 18 % of all samples (31 % of samples with a uniform police presence). There was no significant difference in the number of police in partial riot gear between normal and increased risk situations ($\chi^2 = 0.36, p > .05$). In 16 % of samples (28 % of samples with uniform police presence), police riot vehicles were visible. There was no significant difference in the visibility of police riot vehicles between normal and increased risk situations ($\chi^2 = 0.001, p > .05$). During Euro2000, police riot

vehicles were visible in 50 % of samples and significantly more often in increased as compared to normal risk situations (Adang & Cuvelier, 2001).

Contrary to our expectations, positive interpersonal interactions (contacts) between police and fans were noted in only 5 % of all samples (9 % of samples with uniformed police presence), with a non-significant trend towards slightly more contacts in increased risk situations, 7 % versus 10 % ($\chi^2 = 0.76, p > .05$). During Euro2000, positive contacts between fans and police were significantly higher ($\chi^2 = 509.5, p < .001$), noted in 40 % of samples, with a non-significant trend in the opposite direction such that there were fewer contacts in increased risk situations (Adang & Cuvelier, 2001).

Consistent with arrest figures, observations record an incident of disorder as occurring in just 0.4 % of all samples. Moreover, of the few incidents that did occur, all were rated as small by our observers. During Euro2000, 10 % of samples recorded incidents and these were equally divided over small, medium, and large. This difference is highly significant ($\chi^2 = 152.83, p < .001$). A summary of the differences in patterns of police deployment between Euro2000 and Euro2004 is presented in Table 1.

Analysis of Fans' Perceptions

Pre-Tournament Measures

Identification and differentiation. First, overall levels of ingroup identification, similarity, and difference from relevant outgroups were tested. The mean values for ingroup identification, $t(38) = -8.65, p < .001, d = 1.4$, differentiation from hooligans, $t(38) = -7.40, p < .001, d = 1.2$, and perceived similarity with opposition fans, $t(38) = -2.45, p = .019, d = 0.4$, were all significantly above the mid-point value of three. The mean value for perceived similarity with the Portuguese police did not significantly differ from this mid-point, $t(38) = -0.662, p = .51$. This indicates that prior to the tournament fans reported strong ingroup identification, strong differentiation from hooligans, perceived similarity with opposition fans, and relative ambivalence about their similarity with the police. The means and standard deviations for these items are reported in Table 2.

Expectations of intergroup relations in Portugal. The mean values for expectations about intergroup relations with the local population, opposition fans, and the police are displayed in Table 3. Only expectations about intergroup

Table 1
Structured Observational Data Related to Police Deployment at Euro2004 and Euro2000

	Euro 2000	Euro 2004
Total N of samples	664	1896
Samples with uniformed Police	87%	56%
Avg. uniform officers/100 fans	6-50	4-8
Samples with 'riot' Police	15%	0%
Samples with 'riot vehicles'	50%	16%
Samples positive interaction	40%	5%
Samples recording incidents	15%	0.4%

Table 2
Means and Standard Deviations for Pre- and Post-Tournament Measures of Identification and Differentiation

Questionnaire item	Pre-tournament		Post-tournament	
	Mean	SD	Mean	SD
Ingroup identification	3.9	0.62	3.94	0.63
Differentiation from hooligans	4.31	1.1	4.21	1.16
Similarity with opposition fans	3.41	1.04	3.66	0.96
Similarity with PSP	2.9	0.97	2.73	0.88

relations with the local population differed significantly from the mid-point, $t(38) = 4.90, p < .001, d = 0.8$. Thus, whilst fans tended to be positive in their expectations about the local population, they remained relatively unsure about their expected intergroup relationship with the Portuguese police and with opposition fans.

Associations among intergroup expectations, similarity, and difference. The associations among pre-tournament measures were then explored by calculating bivariate correlations. The correlation coefficients are reported in Table 4. Most notably, ingroup identification shows a strong and significant *negative* correlation with fans perceived similarity with the Portuguese police, $r = -.498, p = .001$. Moreover, perceived similarity with opposition fans was positively correlated with expectations regarding intergroup relations with those fans ($r = .460, p = .004$) and with the local population in match cities ($r = .472, p = .002$).

Post-Tournament Measures

Identification and differentiation. As with the pre-tournament measures, post-tournament measures of ingroup identification, $t(60) = 11.68, p < .001, d = 1.5$, differentiation from hooligans, $t(62) = -8.21, p < .001, d = 1.0$, and perceived similarity to opposition fans, $t(61) = -5.44, p < .001, d = 0.7$, were significantly above the mid-point. Mean perceived similarity to the Police was significantly below the mid-point, $t(58) = -2.66, p = .010, d = 0.3$. Means and standard deviations for these scales are displayed in Table 2.

Table 3
Means and Standard Deviations for Pre- and Post-Tournament Measures of Expectations of Intergroup Relations in Portugal

Questionnaire item	Pre-tournament		Post-tournament	
	Mean	SD	Mean	SD
Experience of ingroup	n.a.	n.a.	4.00	0.81
Intergroup relations with local population	3.59	0.75	3.84	0.84
Intergroup relations with opposition fans	3.08	0.88	3.84	0.78
Intergroup relations with the PSP	3.09	0.92	3.68	0.88

Table 4
Pre- and Post Tournament Correlations Between Measures of Intergroup Relations, Identification, and Differentiation

Item	Ingroup identification	Differentiation from hooligans	Similarity with opposition fans	Similarity with PSP	Relations with population	Relations with opposition fans	Relations with PSP
Ingroup Identification		<i>0.051</i>	<i>-0.009</i>	<i>0.421**</i>	<i>0.066</i>	<i>-0.113</i>	<i>0.15</i>
Differentiation from Hooligans	<i>-0.009</i>		<i>0.454**</i>	<i>0.147</i>	<i>0.248~</i>	<i>0.208</i>	<i>0.152</i>
Similarity with Opposition fans	<i>-0.149</i>	<i>0.23</i>		<i>0.296*</i>	<i>0.153</i>	<i>0.14</i>	<i>0.221</i>
Similarity with PSP	<i>-0.498**</i>	<i>0.006</i>	<i>0.199</i>		<i>0.433**</i>	<i>0.128</i>	<i>0.525**</i>
Relations with Population	<i>-0.257</i>	<i>0.283</i>	<i>0.472**</i>	<i>0.502*</i>		<i>0.516**</i>	<i>0.795**</i>
Relations with Opposition fans	<i>-0.179</i>	<i>0.086</i>	<i>0.460**</i>	<i>0.267</i>	<i>0.314</i>		<i>0.429**</i>
Relations with PSP	<i>-0.132</i>	<i>0.103</i>	<i>0.227</i>	<i>0.296</i>	<i>0.655**</i>	<i>0.416**</i>	

Note. ** Correlation is significant at $p < .01$ level (two tailed); * Significant at $p < .05$ level (two tailed); ~ significant at $p = .05$ (two tailed). Post tournament correlations are in italic font and are located in the top right portion of the table.

Experiences of ingroup behaviour and intergroup relations with the police. As with the pre-tournament expectation measure, fans mean experience of intergroup relations with the local population, $t(62) = 10.33, p < .001, d = 1.3$, was significantly above the mid-point (i.e., positive). In addition, the post-tournament means for experience of intergroup relations with opposition fans, $t(62) = 7.98, p < .001, d = 1.0$, and experience of intergroup relationships with the Portuguese police, $t(60) = 6.04, p < .001, d = 0.8$, were significantly above the mid-point. Finally, the post-tournament measure of experience of the ingroup, $t(56) = 9.22, p < .001, d = 1.2$, was also significantly above the mid-point. The means and standard deviations for these measures are displayed in Table 3.

Comparison of pre-and post tournament measures. Differences between pre- and post-tournament items were tested using between-subjects t -tests. There were no significant differences in levels of ingroup identification, differentiation from hooligans, similarity with opposition fans, or similarity to the police (all t 's < 1).

In order to compare pre- and post-tournament measures relating to intergroup relations, only those items present in both the pre-and post-tournament questionnaires were compared. Thus, scales relating to opposition fans, local population, and policing consisted of one, two, and three items respectively. Post-tournament scores were significantly more positive on all three of these intergroup relations measures: local population, $t(100) = -2.63, p = .01, d = 0.5$; opposition fans, $t(99) = -4.35, p < .001, d = 0.9$; and the Portuguese police, $t(99) = -4.14, p < .001, d = 0.8$.

Post-tournament associations among intergroup experiences, similarity, and difference. Bivariate correlations among the post-tournament measures were calculated. These coefficients are reported in Table 4. The expected positive correlation between intergroup relations and perceived similarity with opposition fans did not emerge. Rather, the measures of intergroup relations in match cities were significantly and positively correlated with perceived similarity to the Police. Moreover, in marked contrast to the pre-tournament measures, ingroup identification was strongly and *positively* correlated with perceived similarity to the Police, $r = .421, p = .001$.

In order to test the reliability of this change in the association between ingroup identification and similarity with the police, an interaction analysis was performed using a general linear model (ANOVA), with similarity with the police as the outcome variable and ingroup identification, time (pre- vs. post-tournament), and the ingroup identification x time interaction as predictors. As expected, the interaction between ingroup identification and time was highly significant, $F(1, 94) = 25.39, p < .001, \eta_p^2 = .219$, indicating that the association between ingroup identification and similarity with the police was significantly more positive after the tournament than it was before.

Discussion

This paper set out to present an integrated quantitative analysis of both policing practice by the PSP during Euro2004 and of England fans' collective perceptions and behaviour. The structured observational data confirm a low profile orientation of PSP deployment on the one hand and low levels of disorder

on the other. Moreover, the levels of visible uniform police, riot police, riot vehicles, and incidents were all significantly lower than Euro2000. In this respect, the data from the structured observations suggest that where crowds were gathered (even during events defined by the authorities as posing increased risk) there was often no obvious uniformed police presence. Moreover, when uniform police were present, they were in normal uniform and in small numbers. Perhaps most remarkably of all, the data suggest that the visibility of paramilitary style police was virtually zero and that there were no major incidents of disorder in match cities.

The survey data suggest that fans experienced this form of policing and their intergroup relationships with other fans in match cities as legitimate, more so than they expected prior to the tournament. In this context of the subjective legitimacy of intergroup relationships, there is also evidence that England fans maintained bonds of common social identity with fans of other nations and a strong psychological differentiation from hooligans. Moreover, these measures of differentiation and similarity were, as predicted by ESIM, positively correlated with expectations of legitimacy in fans intergroup relations with other groups. It is most striking to note that the data suggest that the experience of legitimate policing changed the association between ingroup identification and perceived similarity, or identification with, the police. Prior to the tournament strong identification as an England fan implied *dissimilarity* to police in match cities whereas following the tournament it implied *similarity*.

Limitations

The data therefore provide support for the theoretical analysis set out by Stott et al. (2007). This contribution notwithstanding, it is important to address potential limitations of the current data. First, given the relatively small sample of fans responding to our survey and the lack of control we had over our participants, there are issues relating to the reliability and validity of the questionnaire data. Second, comparable quantitative data from a context of high-profile policing during the tournament would have allowed for direct comparisons, but such data were not available. Third, it was not possible to control for any effect of the legislation, alluded to in the introduction, which prevented large numbers of alleged hooligans from attending the tournament.

Gathering any data on the processes governing collective disorder at international football tournaments is in and of itself politically sensitive, highly opportunistic, physically and psychologically demanding, and potentially dangerous. It is therefore difficult to achieve the same levels of control that one might expect from studying social psychological processes in more controlled settings (e.g., University laboratories using undergraduate students). Nonetheless, we hold the view that it is absolutely critical for psychologists to contribute to the study of crowd events and necessary to gather data from the field if we are to fully explore the explanatory power and validity of our theoretical models. In this respect, field-based data should not be rejected simply because of its inherent weaknesses and the priority many within the discipline place upon epistemological certainty. In this respect, our view is that the extensive quantitative data presented here are

unprecedented, actually remarkably thorough, and a valuable contribution to the wider body of scientific research on crowd psychology and behaviour.

Turning specifically to the problems associated with the survey data, there is no a-priori reason to assume that the fans involved in our study were somehow a particularly unrepresentative sample of England fans or were not present at the tournament. The vast majority of the sample explicitly claimed to have been highly committed and experienced fans who regularly attend England games across Europe on an ongoing basis. It would therefore be highly unusual for them, given this level of dedication, not to attend a major international tournament involving the England team. Indeed, one of the primary reasons for fans travelling to many fixtures abroad is that it provides them with “caps” or priority access to the very limited supply of match tickets for the finals when, and if, England eventually qualify. Moreover, one of the defining characteristics of England fans is loyalty, which itself is defined in terms of travelling to England fixtures whenever possible. Moreover, almost all our respondents had a good deal to say in the open ended sections about their experiences of policing in Portugal as this contrasted with their experiences elsewhere (the data from which is reported in more detail in Stott et al., 2007; Stott & Pearson, 2007).

As we have made clear the sample of fans in this study is relatively small and we cannot be certain about the number of participants who completed both the pre- and post test measures. Whilst we accept that these are important issues, our data indicates that a perception of legitimacy in intergroup relationships with the police in match cities was present among our sample of highly identified category members. It also demonstrates that a transformation in perceived similarity to the police in match cities did occur among our sample between the beginning and the end of the tournament. It is therefore necessary to explain how such a change has taken place amongst our sample and our argument is that these fans perception of the legitimacy of the policing of the tournament was central to this transformation and that such data provides a small but important empirical contribution to a theoretical explanation of the absence of disorder in match cities at the tournament itself.

The observational data reported in this paper are also consistent with a data set, including official police reports and semi-structured observations, gathered as part of the wider project (see Stott et al., 2007; Stott & Pearson, 2007, for further details). This multiple method approach allows any sampling biases in the observational data presented here to be detected through a process of triangulation among multiple data sources gathered using different methods (Denzin, 1989). For example, the fact that there was only one England fan recorded in official police statistics as arrested for a violence related offence in match cities corresponds with police records that no officers in the *Corpo de Intervenção* officially drew their batons throughout the entire tournament. This data, in turn, provides validation of the findings emerging from the structured observations. Thus, while sampling bias may conceivably be an issue in some aspects of the present analysis, it is most unlikely to have influenced the key findings regarding overall policing profile and levels of public order.

In terms of the relative absence of systematic data from high-profile policing contexts, this is clearly true in the present paper. Indeed, the fact that we are not able to quantifiably compare different forms of policing with variations in the

levels of conflict largely reflects the success of the PSP in implementing the low profile approach throughout Portugal (i.e., all data sources suggest that there were very few and only short deployments of riot police and certainly none at all in our samples) and the overall absence of major incidents of disorder in match cities. However, Stott et al. (2007) do provide a relatively detailed qualitative account of two major incidents of disorder during the tournament. Crucially, though, these did not occur in a match city but in Albufeira, a small resort town in Portugals Algarve region controlled by Portugals second major police force, the Guarda Nacional Republicana (GNR). Stott et al. (2007) have provided evidence that not only did the GNR utilise forms of high profile policing but also that the rioting was associated with forms of collective psychology and dynamics predicted by the ESIM (e.g., Stott et al., 2001; Stott & Reicher, 1998a).

Finally, it is in some sense obvious that reducing the number of hooligans (i.e., those intent of creating disorder) at an international tournament is likely to impact upon the tendencies for incidents to develop. However, it has been argued elsewhere that, despite the extensive policing and legislative measures that were in place in the U.K., English hooligans were actually present during Euro2004 and did try to initiate disorder in match cities (see Stott et al., 2007). Thus, the arguments we make here are concerned not so much with the presence or absence of hooligans or with relatively minor and isolated acts of hooliganism. Instead, our analysis is focused upon the social and psychological conditions that transform these acts of hooliganism into the widespread collective rioting that has been witnessed at almost every other international soccer tournament held in continental Europe since the 1980s (for a historical overview, see Stott & Pearson, 2007). Our contention is that major riots were not witnessed in match cities during Euro2004 because the low profile adopted by the PSP undermined the ability of hooligans to draw previously non-hooligan fans into conflict during crowd events. Moreover, should more indiscriminate forms of policing have been adopted then our contention is that major riots would have developed in match cities precisely because (a) hooligan fans would have been more influential in the crowd, and (b) non-hooligan fans would have been drawn into the dynamics of conflict. While keeping in mind the role of other factors, such as the activities of fan organisations or the efforts of local Government agencies (e.g., Miles, 2000; Perryman, 2006), this suggests a key role for policing profile in generating public order outcomes, over and above (but not in isolation from) the mere presence or absence of hooligans.

Implications: Theory and Practice in Public Order Policing

Turning to the strengths of the present research, it represents as complete and comprehensive a study of crowd behaviour as yet exists in the literature. This is in no small part because of the ability of the current study to directly relate police practices to collective (fan) perceptions and behaviour within the same event using quantifiable data. In so doing, it has come closer than any other study to presenting the ideal, fully interactive analysis of intergroup dynamics called for in the broader literature on crowd psychology (e.g., Reicher, 1996; Stott & Reicher, 1998b).

The current study therefore provides an important body of evidence in support of the ESIM and is at the same time consistent with other analyses of low profile or negotiated management styles of public order management (e.g., Adang & Cuvelier, 2001; De Lint, 2005, p. 196; King & Waddington, 2004, 2005; Waddington & King, 2005). Peter Waddington has consistently argued in favour of the use of paramilitary policing as a means of assisting in the maintenance of public order during crowd events (Waddington, 1987, 1991, 1993, 1994). His contention is that the discipline, combined with the strong command and control of paramilitary policing is effective at undermining tendencies that officers have to invoke disorder through uncoordinated and indiscriminate discretionary actions. Crucially, he also argues that paramilitary interventions against the crowd must be underpinned by information, such that the police are able to accurately target the use of force. In contrast, Jefferson (1987, 1990) argues that paramilitary policing carries inherent dangers of precipitating and exacerbating collective violence. He proposes a four stage process where police expectations of disorder can lead inevitably to disproportionate and indiscriminate use of force by paramilitary police, a process through which widespread disorder emerges as a form of self fulfilling prophesy (see also Stott & Drury, 2000; Stott & Reicher, 1998b).

The data from this study tends to be more consistent with Jefferson's position than that of Peter Waddington's. However, data gathered from ethnographic observations of Euro2004 reported in Stott et al. (2007) points toward the important role played by squads of non-uniformed officers who were operating within crowds gathered in match cities. That study also shows that large squads of riot police were present but being kept deliberately out of sight ready for rapid deployment. These forms of police activity would not have been recorded in the data obtained by the structured observations reported in the current paper. Stott et al. (2007) also point out that the non-uniformed officers were able to gather real time information on and react quickly and sensitively to any emerging sources of risk to public order. Thus, whilst the low profile approach of the PSP may have avoided the obvious use of paramilitary police, it still retained the capability to utilise them. Thus, it was the early and proactive deployment of non-paramilitary police that appears to have increased the capacity for the PSP to engage the type of information led targeted use of force central to Peter Waddington's model. However, the current study suggests that it was the deployment of non-paramilitary tactics to locate and deal with emergent problems in the first instance that meant that public order was maintained and that the paramilitary police were hardly, if ever, required.

This study also advances upon the existing literature by providing a unique test of the relationships between low profile policing and the social psychology of crowd events at an international football tournament. It is more specific to note that it supports the contention that the efficiency and effectiveness of low profile policing lies in its ability to manage risk to public order during crowd events in ways that generate (or at least maintain) shared perceptions among crowd participants of the legitimacy of their intergroup relationships with the police (Adang & Stott, 2004; Stott & Adang, 2003a,b; Stott et al., 2007). In other words, our data suggests that low profile policing works because it manages crowd events in such a manner that avoids the forms of intergroup interaction, collective psychology and intragroup relations that the literature on crowd psychology proposes are

necessary for widespread escalation and rioting to occur (e.g., Drury & Reicher, 2000; Reicher, 1984, 1996; Stott et al., 2007; Stott & Drury, 2000; Stott et al., 2001; Stott & Reicher, 1998a, 2003b).

The current study also advances that literature by suggesting the importance of identification between crowd participants and the police, which may function as the psychological tool through which public order can be successfully maintained. Put in slightly different terms, the study suggests that after experiencing legitimacy in intergroup relations with the PSP those who identified strongly as England fans began to psychologically identify with the police. This also suggests that those England fans that saw themselves as different from the police began to become psychologically disassociated from their social category. This finding echoes Stott et al.'s (2007) contention that the low profile policing context affects the extent to which those seeking violence become marginalised and disempowered within the wider community of fans, possibly over the longer term (see also Stott et al., 2001).

The current paper has therefore, despite its limitations, made an important empirical and theoretical contribution to understanding the relationship between police profile, fan psychology, and levels of disorder in the context of football matches with an international dimension. It demonstrates that an explanation of the absence of disorder in terms of the social psychology of crowd dynamics is at the very least possible, and, we would suggest, parsimonious and powerful. This study therefore highlights the potential policy benefits of the ESIM approach and the importance of ensuring that similar research on the policing of international tournaments can and does take place in the future. In particular, it is important for future research to address the limitations of the current study. In particular, it would be preferable to develop ways of recruiting much larger numbers of fans before the tournament begins and track them across, during, and at various times following their participation in crowd events. This would provide more reliable longitudinal data on the impact of the social context upon their involvement in collective action and their underlying psychology. Moreover, it would be preferable to gain data from and access to the many different police forces and other authorities involved in policing fans across multiple international tournaments, in order that stronger comparisons between different approaches can be achieved. Of course, gathering such high quality data in the context of a major international football tournament would be a major achievement and we believe one that would only be possible with international political support at the highest levels. In this respect, recent initiatives to develop a European wide training and research network, put forward at a "High Level Conference Towards and EU Strategy Against Violence in Sport" – collaboratively organized by the European Commission, the Portuguese Presidency of the Council, the European Parliament, and UEFA in Brussels on the 28th and 29th November, 2007 – bode well for establishing the necessary framework for such future research to even be possible.

While the evidence from this paper has been derived from a football tournament in the European context, it demonstrates the utility of the study of football crowds for advancing theoretical understanding of crowd dynamics, social identity processes, and public order policing. Consequently, the dynamics examined here have clear relevance to the management and theoretical understanding of crowds in different contexts (e.g., celebratory crowds at New Year's Eve in

celebrations in Trafalgar Square, London, or in Times Square, New York; festival crowds, such as the Notting Hill Carnival in London or the Mardis Gras in New Orleans; other sporting crowds, such as those attending the Olympics in London or elsewhere). One of the central issues this research raises then is the relationship between psychological theory, policy, and practice. Here and elsewhere, it has been demonstrated that productive relationships can be developed between scientific theories on the one hand and police operational practice on the other (Adang & Cuvelier, 2001; Stott et al., 2007; Stott & Pearson, 2007; Waddington, 2005). Indeed, subsequent to our research on Euro2004, the European Council Working Party on International Police Cooperation (2006) adopted recommendations concerning the dynamics of risk to public order and their relationship to police tactical profile suggested by the pre-tournament model of good practice, the current data and other related research (Adang & Cuvelier, 2001; Adang & Stott, 2004; Stott & Adang, 2003a,b; Stott et al., 2007).

Despite these policy developments, the policing of football in Europe can still be based upon the rather outdated theoretical idea that violence is an inherent characteristic of the crowd (cf. Allport, 1924; Le Bon, 1895/1947; Zimbardo, 1969). The police consequently see their role as simply to react to, control, contain, and disperse the inherent tendencies of the crowd toward disorder and to undermine the actions of violent minorities seeking to hijack the irrational mass (Drury et al., 2003; Stott & Reicher, 1998b). In contrast, this study adds further to the now substantial body of evidence that supports the ESIM proposition that major riots can be the outcome of patterns of intergroup interaction that occur *during* a crowd event. As such, police actions can be understood as central to the *development* of disorder, as well as being a response to it (e.g., Stott & Drury, 2000). In line with other social identity-based models of social influence (Turner, 1991), this suggests that the ability of a violent group to shape the norms of a crowd can be dependent upon the surrounding intergroup context, of which police actions are invariably an integral component. This suggests that the ability of violent groups to provoke a riot can ironically be enhanced by deterrence, zero tolerance, escalated force or what Vitale (2005) refers to as “command and control” policing. These styles of policing rely on the use of large numbers of paramilitary police to intensively regulate the movements and behaviours of the crowd. Such policing may, rather ironically given the underlying intentions of conflict reduction, actually generate the dynamics of subjective illegitimacy and empowerment through which riots can and do emerge. Put slightly differently, command and control policing may paradoxically provide the intergroup context through which violent groups can become *more* influential in the crowd than they otherwise might have been.

This current study demonstrates that, where policing is in line with the principles of negotiated management, there are associated and measurable reductions of violence among those social categories consistently and ubiquitously associated with disorder during previous and indeed subsequent events. Moreover, our data also suggest that the group-level social psychological processes associated with negotiated management could also function to positively transform previously negative relationships between high risk groups and the police and lead to the psychological marginalization of those who seek violence. This finding suggests that while high profile, coercive, and ultimately militaristic

policing (i.e., baton charges, water cannon, tear gas) may well be effective at gaining short term control of a physical location, such a heavy hand could actually function to generate and maintain a negative relationship over the long term between the police and specific communities. This may not only sustain conflictual relations between police and that community but also may also create the conditions under which that community supports and aids these violent groups. Thus, the control and containment of groups suspected of seeking to initiate premeditated conflict may well have an impact. However, the wider research on football crowds suggests that such a policy certainly does not prevent major incidents from developing. Moreover, such strategies involve significant costs both financial and civil libertarian. We should therefore always be mindful in democratic societies of developing alternative strategies of conflict management that are based upon scientific evidence, are less confrontational, more efficient, and are less costly in terms of civil liberties and human rights.

References

- Adang, O. M. J. (1988). Voetbalvandalisme geobserveerd. *Tijdschrift voor Criminologie*, 4, 302–312.
- Adang, O. M. J. (1991). Gedrag van mensen tijdens rellen. In P. B. Defares & J. D. van der Ploeg (Eds.) *Agressie*. (pp.75–87). Assen, The Netherlands: Van Gorcum.
- Adang, O. M. J. (1998). *Hooligans, autonomen, agenten. Geweld en politie-optreden in relsituaties*. Alphen aan den Rijn, The Netherlands: Samson.
- Adang, O. M. J., & Cuvelier, C. (2001). *Policing Euro2000: International police co-operation, information management and police deployment*. Tandem-Felix: Ubbergen.
- Adang, O. M. J., & Stott, C. J. (2004). *Preparing for Euro2004: Policing international football matches in Portugal*. Unpublished report for the Portuguese Public Security Police.
- Allport, F. H. (1924). *Social psychology*. Boston: Houghton Mifflin Co.
- Armstrong, G. (1998). *Football hooligans: Knowing the score*. Oxford, England: Berg.
- Baker, D. (2005). *Batons and blockades: Policing industrial disputes in Australasia*. Beaconsfield, Victoria, Australia: Melbourne Publishing.
- Brug, H. van der (1989). *Voetbalvandalisme in Nederland*. Amsterdam: Stichting Het Persinstut.
- Crabbe, T. (2003). The public gets what the public wants: England football fans, 'truth' claims and mediated realities. *International Review for the Sociology of Sport*, 38, 413–425.
- Dal Lago, A. (1990). *Descrizione di una Battaglia: I rituali del Calcio*. Bologna: Il Mulino.
- De Lint, W. (2005). Public order policing: A tough act to follow. *International Journal of the Sociology of Law*, 31, 179–199.
- Della Porta, D., & Reiter, H. (Eds.). (1998). *Policing protest: The control of mass demonstrations in western democracies*. Minneapolis, MN: University of Minnesota Press.
- Denzin, N. K. (1989). *The research act: A theoretical introduction to sociological methods* (3rd ed.). Englewood. Cliffs, NJ: Prentice Hall.
- Doosje, B., Ellemers, N., & Spears, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of Experimental Social Psychology*, 31, 410–436.
- Drury, J., & Reicher, S. (1999). The intergroup dynamics of collective empowerment:

- Substantiating the social identity model of crowd behaviour. *Group Processes and Intergroup Relations*, 2, 1–22.
- Drury, J., & Reicher, S. D. (2000). Collective action and psychological change: The emergence of new social identities. *British Journal of Social Psychology*, 39, 579–604.
- Drury, J., Stott, C. J., & Farsides, T. (2003). The role of police perceptions and practices in the development of public disorder. *Journal of Applied Social Psychology*, 33, 1480–1500.
- Dunbar, R. I. M. (1976). Some aspects of research design and their implications in the observational study of behaviour. *Behaviour*, 58(1–2): 79–98.
- Dunning, E. (2000). Towards a sociological understanding of football hooliganism as a world phenomenon. *European Journal of Criminal Policy and Research*, 8, 141–62.
- Dunning, E., Murphy, P., & Williams, J. (1988). *The roots of football hooliganism: An historical and sociological study*. London: Routledge & Kegan Paul.
- Frosdick, S., and Marsh, P. (2005). *Football hooliganism*. Cullompton, United Kingdom: Willan.
- Hall, A., & De Lint, W. (2003). Policing labour in Canada. *Policing & Society*, 13, 219–234.
- Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social Identity Salience and the Emergence of Stereotype Consensus. *Personality and Social Psychology Bulletin*, 25, 809–818.
- Home Office. (2005). *Impact of measures introduced by the Football (Disorder) Act 2000 Report to Parliament*. U. K. Home Office. Retrieved August 13, 2008, from <http://www.homeoffice.gov.uk/documents/Football-Disorder-2006.pdf?view=Binary>
- Hutt, S. J., & Hutt, C. (1974). *Direct observation and measurement of behavior*. Springfield, IL: Charles C. Thomas.
- Independent Football Commission. (2004). *Report on Euro2004*. Retrieved August 13, 2008, from <http://www.theifc.co.uk>
- Innes, M. (2005). Why soft policing is hard: On the curious development of reassurance policing, how it became neighbourhood policing and what this signifies about the politics of police reform. *Journal of Community and Applied Social Psychology*, 15, 156–169.
- Jefferson, T. (1987). Beyond paramilitarism. *British Journal of Criminology*, 27, 47–53.
- Jefferson, T. (1990). *The case against paramilitary policing*. Milton Keynes, United Kingdom: Open University Press.
- Kerr, J. (1994). *Understanding soccer hooliganism*. Buckingham, England: Open University Press.
- King, A. (1995). Outline of a practical theory of football violence. *Sociology*, 29, 635–651.
- King, M., & Waddington, D. (2004). Coping with disorder? The changing relationship between police public order strategy and practice – A critical analysis of the Burnley Riot. *Policing and Society*, 14, 118–137.
- King, M., & Waddington, D. (2005). Flashpoints revisited: A critical application to the policing of anti-globalisation protest. *Policing & Society*, 15, 255–282.
- Le Bon, G. (1947). *The crowd: A study of the popular mind*. London: Ernest Benn. (Original work published 1895)
- Lehner, P. N. (1979). *Handbook of ethological methods*. New York: Garland STPM Press.
- Marsh, P., Rosser, E., & Harre, R. (1978). *The rules of disorder*. London: Routledge & Keegan Paul.
- McPhail, C. (1991). *The myth of the madding crowd*. New York: Aldine de Gruyter.
- Miles, K. (2000). *An English fan abroad*. Edinburgh, United Kingdom: Canongate.
- Perryman, M (2006). *Ingerland travels with a football nation*. London: Simon & Schuster.

- Pilz, G. (1996). Social factors influencing sport and violence: On the problems of football hooliganism in Germany. *International Review for the Sociology of Sport*, 31, 49–68.
- Police Cooperation Working Party of the Council of the European Union. (2006). European Union handbook on international police cooperation and measures to prevent and control violence and disturbances in connection with football matches with an international dimension. *Official Journal of the European Union (English Edition)*, 49, C32. Retrieved August 13, 2008, from <http://europa.eu.int/eur-lex/lex/JOIndex.do?year=2006&serie=C&textfield2=322&Submit=Search>
- Popplewell, O. (1986). *Committee of Enquiry into Crowd Safety and Control at Sports Grounds*. Final Report. London: HMSO, 6.
- Poulton, E. (2005). English media representation of football related disorder: “Brutal, shorthanded and simplifying”? *Sport and Society*, 8, 27–47.
- Reicher, S., Stott, C., Cronin, P., & Adang, O. (2004). An integrated approach to crowd psychology and public order policing. *Policing: An International Journal of Police Strategies and Management*, 27, 558–572.
- Reicher, S., Stott, C., Drury, J., Adang, O., Cronin, P., & Livingstone, A. (2007). Knowledge based public order policing: Principles and practice. *Policing: A Journal of Policy and Practice*, 1, 403–415.
- Reicher, S. D. (1984). The St. Pauls riot: An explanation of the limits of crowd action in terms of a social identity model. *European Journal of Social Psychology*, 14, 1–21.
- Reicher, S. D. (1996). “The Battle of Westminster”: Developing the social identity model of crowd behaviour in order to explain the initiation and development of collective conflict. *European Journal of Social Psychology*, 26, 115–134.
- Roversi, A. (1991). Football violence in Italy. *International Review for the Sociology of Sport*, 26, 311–332.
- Russell, G., & Goldstein, J. (1995). Personality differences between Dutch football fans and non-fans. *Social Behaviour and Personality*, 23, 199–204.
- Sheptycki, J. (2005). Policing political protest when politics go global: Comparing public order policing in Canada and Bolivia. *Policing & Society*, 15, 327–352.
- Slater, P. J. B. (1978). Data collection. In P. N. Colgan (Ed.), *Quantitative ethology* (pp.7–24). New York: John Wiley & Sons.
- Stott, C. J., & Adang, O. M. J. (2003a) *Policing Football Matches with an International Dimension in the European Union: Understanding and managing risk*. Retrieved August 13, 2008, from <http://www.liv.ac.uk/psychology/staff/cstott.html>
- Stott, C. J., & Adang, O. M. J. (2003b). *Crowd psychology and public order policing*. Paper presented at PSP Euro2004 Strategy meeting, Instituto Superior de Ciências Policiais e Segurança Interna, Lisbon, Portugal. Retrieved August 13, 2008, from <http://www.liv.ac.uk/psychology/staff/cstott.html>
- Stott, C. J., Adang, O. M. J., Livingstone, A., & Schreiber, M. (2007). Variability in the collective behaviour of England fans at Euro2004: Hooliganism, public order policing and social change. *European Journal of Social Psychology*, 37, 75–100.
- Stott, C. J., & Drury, J. (2000). Crowds, context and identity: Dynamic categorization processes in the Poll Tax Riot. *Human Relations*, 53, 247–273.
- Stott, C. J., Hutchison, P., & Drury, J. (2001). Hooligans abroad? Inter-group dynamics, social identity and participation in collective disorder at the 1998 World Cup Finals. *British Journal of Social Psychology*, 40, 359–384.
- Stott, C. J., Livingstone, A., & Hoggett, J. (2008). Policing football crowds in England and Wales: A model of good practice? *Policing and Society*, 18, 1–24.
- Stott, C., & Pearson, G. (2006). Football banning orders, proportionality, and public order policing. *The Howard Journal*, 45, 241–254.
- Stott, C. J., & Pearson, G. (2007). *Football hooliganism, policing and the war on the “English Disease.”* London: Pennant Books.

- Stott, C. J., & Reicher, S. D. (1998a). How conflict escalates: The inter-group dynamics of collective football crowd violence. *Sociology*, *32*, 353–377.
- Stott, C. J., & Reicher, S. D. (1998b). Crowd action as inter-group process: Introducing the police perspective. *European Journal of Social Psychology*, *26*, 509–529.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations*. Monterey, CA: Brooks Cole.
- Turner, J. C. (1991). *Social influence*. Milton Keynes, England: Open University Press.
- Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454–463.
- Turner, R. H., & Killian, L. M. (1987). *Collective behavior*. (3rd edition). Englewood Cliffs, NJ: Prentice Hall.
- Vitale, A. (2005). From negotiated management to command and control: How the New York Police Department polices protests. *Policing & Society*, *15*, 283–304.
- Waddington, D., & King, M. (2005). The disorderly crowd: From classical psychological reductionism to socio-contextual theory – The impact on public order policing strategies. *The Howard Journal*, *44*, 490–503.
- Waddington, P. A. J. (1987). Towards paramilitarism? Dilemmas in the policing of public order. *British Journal of Criminology*, *27*, 37–46.
- Waddington, P. A. J. (1991). *The strong arm of the law: Armed and public order policing*. Oxford, England: Clarendon.
- Waddington, P. A. J. (1993). The case against paramilitary policing considered. *British Journal of Criminology*, *33*, 353–373.
- Waddington, P. A. J. (1994). Coercion and accommodation: Policing public order and the public order act. *British Journal of Sociology*, *45*, 367–385.
- Waddington, D. P. (2005). *Policing public disorder: Theory and practice*. Cullompton, United Kingdom: Willan.
- Williams, J., Dunning, E., & Murphy, P. (1989). *Hooligans abroad*. (2nd edition). London: Routledge.
- Zimbardo, P. G. (1969). The human choice: Individuation, reason and order versus de-individuation, impulse and chaos. In W. J. Arnold and D. Levine (Eds.), *Nebraska symposium on motivation* (Vol. 17). Lincoln, NE: University of Nebraska Press.

Appendix A

Breakdown of Categories Used in the Structured Observations

Location
 Time
 Observer stationary or mobile
 Number persons present in total (exclusive of police)
 Team 1 (country code)
 Team 2 (country code)
 Fans of other country present? (no or country code)
 Number fans team 1

Number fans team 2
 Number of media representatives
 Number non-police security
 Number unknown civilian/ others
 Number police in total
 Number of plain clothes police
 Number of standard uniform police
 Number of partial riot police
 Number of full riot police
 Number normal police vehicles
 Number of police vans
 Number water cannon
 Number police horses
 Number police dogs
 Number fans team 1 agitated (not calm)
 Number fans team 1 shouting/ singing (not quiet)
 Number fans team 2 agitated (not calm)
 Number fans team 2 shouting/ singing (not quiet)
 Number of team 1 fans within touching distance of team 2 fans
 Number of team 2 fans within touching distance of team 1 fans
 Number of police within touching distance to fans
 Number of fans team 1 in positive interaction with team 2 fans
 Number of fans team 2 in positive interaction with team 1 fans
 Number of fans team 1 in negative interaction with team 2 fans
 Number of fans team 1 in negative interaction with team 2 fans
 Number of police in positive interaction with team 1 fans
 Number of police in negative interaction with team 1 fans
 Number of police in positive interaction with team 2 fans
 Number of police in negative interaction with team 2 fans
 Incident in last 15 min (none, small, medium, large)
 Police intervention in last 15 min (none, stopping, arrest, dispersion)
 Use of force by police in last 15 min (yes/no)

Appendix B

Open Ended Items on Web Based Questionnaire

We would also very much like to hear from you in your own words about your experiences as an England fan during your time in this host city. If you are willing to provide us with your views then please type your responses to the questions in the relevant text box below. Please provide any opinions you have about issues or events that were relevant to you or other England fans during events surrounding this match.

1. If you have not already done so please describe your views upon the overall quality of the policing or stewarding of England fans.

Received July 28, 2006

Revision received June 25, 2008

Accepted June 26, 2008 ■